Claim Amendment

1. (currently amended) A method for speech processing in a distributed-speech recognition system having a front-end and a back-end for recognizing words from speech signals in a time domain, said method comprising the steps of:

transforming the speech signals in the time domain for obtaining spectrum representation of the speech signals in a frequency domain;

transforming the spectrum representation for obtaining speech features in a cepstral domain; ~~extracting speech features from the speech signals~~, wherein the speech features comprise a speech component and a noise component in ~~contain~~ a speech-to-noise ratio;

normalizing the speech features for providing normalized speech features having a reduced speech-to-noise ratio;

filtering the normalized speech features in a frequency domain for reducing the noise component; and

conveying the filtered speech features from the front-end to the back-end.

2. (original) The method of claim 1, wherein the filtering step is carried out with a low-pass filter.

3. (original) The method of claim 1, wherein the filtering step is carried out with a data-driven filter.

4. (original) The method of claim 1, further comprising the step of converting the speech signals from a time domain to a frequency domain prior to extracting the speech features.

5. (original) The method of claim 4, further comprising the step of converting the speech signals to digital signals prior to converting the speech signals from the time domain to the frequency domain.

6. (original) The method of claim 4, wherein the time-to-frequency domain conversion is carried out by a Fast Fourier Transform in order to compute a magnitude spectrum and provide a plurality of magnitude spectrum values.

7. (original) The method of claim 6, further comprising the step of non-linearly modifying the magnitude spectrum in order to generate a plurality of logarithmically-warped magnitude spectrum values.

8. (original) The method of claim 7, further comprising the step of assembling the logarithmically-warped magnitude spectrum values in order to produce a set of feature parameters representative of the speech features.

9. (currently amended) A distributed speech recognition front-end comprising:

first means, responsive to a speech signal in a time domain, for obtaining spectral representation of the speech signal in a frequency domain;

second means, responsive to the spectral representation, for extracting speech features in the cepstral domain ~~from said speech signal~~ and for providing a first signal indicative of the extracted speech features, the extracted speech features comprising a speech component and a noise component in a speech-to-noise ratio;

~~second~~ third means, responsive to the first signal, for normalizing the extracted speech features in order to provide normalized speech features having a reduced speech-to-noise ratio and for providing a second signal indicative of the normalized speech features;

~~third~~ fourth means, responsive to the second signal, for filtering the normalized speech features in [[a]] the frequency domain in order to reduce the noise component in normalized speech features ~~second signal~~ and for providing a third signal indicative of the filtered speech features; and

means for conveying the third signal to a distributed speech recognition back-end in order for the back-end to recognize words representative of the speech signal from the third signal.

10. (original) The front-end of claim 9, wherein the third means comprises a data-driven filter.

11. (original) The front-end of claim 9, wherein the third means comprises a low-pass filter.

12. (original) The front-end of claim 9, wherein the first means comprises:

a time-domain, pre-processing device to convert the speech signal to a digital signal;

a time-to-frequency domain conversion device to provide a set of magnitude spectrum values from the digital signal; and

an assembly device to assemble the set of magnitude spectrum values into the speech features.

13. (original) The front-end of claim 9, wherein the third signal has a sampling rate, said front-end further comprising means to reduce the sampling rate prior to conveying the third signal to the distributed signal recognition back-end.

14. (currently amended) A distributed speech recognition system for processing a speech signal, said system comprising:

a front-end, responsive to the speech signal, for extracting speech features in a cepstral domain from the speech signal and for providing a first signal indicative of the extracted speech features, the extracted speech comprising a speech component and a noise component in a speech-to-noise ratio; and

a back-end, responsive to the first signal, for recognizing words representative of the speech signals and for providing a second signal indicative of the recognized words, wherein

the front-end has means to normalize the extracted-speech features for providing normalized speech features with a reduced speech-to-noise ratio and means to filter the normalized speech features in order to reduce the noise component in the speech signal.

15. (original) The system of claim 14, wherein the filtering means comprises a low-pass frequency filter.

16. (original) The system of claim 14, wherein the filtering means comprises a data-driven filter.

17. (currently amended) A speech recognition feature extractor for extracting speech features from a speech signal, comprising:

a time-to-frequency domain transformer for generating spectral magnitude values in a frequency domain of the speech signal and for providing a first signal indicative of the spectral magnitude values;

a feature generator, responsive to the first signal, for generating a plurality of feature vectors <u>in a cepstral domain</u> and for providing a second signal indicative of the generated feature vectors<u>, the feature vectors comprising a speech component and a noise component in a speech-to-noise ratio</u>;

a normalizing means, responsive to the second signal, for normalizing the generated feature vectors <u>in order to provide normalized feature vectors having a reduced speech-to-noise ratio</u> and for providing a third signal indicative of the normalized feature vectors; and

a frequency filtering means, responsive to the first signal, for reducing <u>the</u> noise <u>component</u> in the normalized feature vectors and for providing the extracted speech features indicative of the noise-reduced feature vectors.

18. (original) The extractor of claim 17, wherein the frequency filtering means comprises a low-pass filter.

19. (original) The extractor of claim 17, wherein the frequency filtering means comprises a data-driven filter.

20. (currently amended) A communication device having a voice input unit to allow a user to input speech signals to the device, and means for providing speech data to an

external apparatus, wherein the external apparatus includes a distributed-speech recognition back-end capable of recognizing speech based on the speech data, said communication device comprising

a front-end unit, responsive to the speech signals, for extracting speech features in a cepstral domain from the speech signals for providing a first signal indicative of the extracted speech features, the extracted speech features comprising a speech component and a noise component in a speech-to-noise ratio, wherein

the front-end includes:

means, responsive to the first signal, for normalizing the extracted-speech features for providing a second signal indicative of the normalized speech features, the normalized speech features having a reduced speech-to-noise ratio, and

means, responsive to the second signal, for filtering the normalized speech features in order to reduce the noise component in ~~the~~ filtered speech features ~~signals~~ and for including the filtered speech features in the speech data.